

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form ([see an example](#)) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Accounting for recent trends in the prevalence of diarrhea in the Democratic Republic of Congo (DRC)
AUTHORS	Kandala, Ngianga-bakwin; Emina, Jacques

VERSION 1 - REVIEW

REVIEWER	Nigel Rollins Scientist World health Organization Switzerland
REVIEW RETURNED	14-Sep-2012

THE STUDY	The statistical methods were not very clear to me nor were the limitations of the MICS and DHS data adequately discussed or considered. For these reasons the investigators were limited in what they could conclude and I did not have confidence in the results. There was also considerable repetition and the overall language of the manuscript could be improved.
RESULTS & CONCLUSIONS	I could not easily understand the data myself in order to come to my own conclusions. Given the limits of the data and the way in which I struggled to assess the results, I did not find the findings and message clear. See uploaded file
GENERAL COMMENTS	General This paper examines the trends in diarrhoeal prevalence among children less than 5 years of age in regions of the Democratic Republic of Congo (DRC) using data from three population based surveys conducted in 1995, 2001 and 2007. The investigators present the background demographics of the Provinces surveyed and comment on the variation between the sample populations. Significant differences in diarrhoeal prevalence are evident over the three surveys. The investigators use different analytical approaches to assess the causes of the changes over time. The primary conclusion was that diarrhoea remains a significant public health problem. The investigators infer that interventions to improve home conditions would reduce diarrhoeal prevalence. While I commend the intention of the investigators to examine the available datasets to understand changes in diarrhoeal prevalence in the DRC, I do not feel the manuscript is strong enough yet to be

	<p>published in the Journal. There is significant repetition and insufficient discussion of the data itself and the limitations of the approaches taken. There are also several grammatical points and style issues that could be improved.</p> <p>Major comments</p> <ol style="list-style-type: none"> 1. The investigators have used data collected in the 1995 and 2001 MIC Surveys and the 2007 DHS to determine relationships between patterns of diarrhoea prevalence in children and causal factors. The analyses and interpretation are therefore fundamentally limited by the quality of the survey data which the investigators obviously cannot change. However, it is a critical limitation which has important consequences for interpretation of the data. While the authors acknowledge this, they do not really examine the implications. For example, in the discussion, it is simply stated that 'there may also be some issues with data quality' but they do not really discuss how these issues might be resolved either with respect to the current analysis (they could corroborate findings by concurrently examining mortality data) or in the future e.g. what it means for sampling in future population surveys or design of questions. 2. Introduction. As much as diarrhoeal burden is important, it would be good to include data on diarrhoeal deaths or all U5 deaths over the same intervals and to comment on the relation between diarrhoea and subsequent health outcomes. The majority of diarrhoeal events are not associated with death and therefore it is important to contextualize the significance of the diarrhoea burden. 3. Methods. 1.2. Data. The investigators comment that the sample designs and questionnaires of the three surveys are reported elsewhere. While the sampling approaches are reasonably understood, the exact questions used in each of the surveys to determine diarrhoeal morbidity is critical to the analysis. It would be relevant to outline the actual wording used in each survey to ascertain information on diarrhoeal prevalence and if/how acute diarrhoea is differentiated from persistent diarrhoea or dysentery. 4. Methods 1.2. Data. The investigators describe how they define high quality water – are similar definitions used in MICS or DHS or are these definitions applied to specific categories of water supply that are included in the survey tools? 5. Table 1. There are major regional variations in the number of children U5 included between Provinces over time. What were the reasons for the change in number of children included? If this is known it would be good to give some details in the methods section. Whatever the reason, the investigators need to comment on the implications. 6. Statistical methods. The description of the decomposition approach is not very clear and leaves the reader (me) unclear about the interpretation. I also found it odd to include a justification of the method as part of the methods section. While it may be a standard approach in demographic analyses, it is not a method that I, as a clinician and public health scientist, have used and feel confident to interpret. The investigators state that the approach is simple and gives a clear understanding of trends. After reading the paper, I cannot say
--	---

	<p>that I clearly understand the trends presented and did not find the application of the decomposition analysis particularly informative.</p> <ol style="list-style-type: none"> 7. Results. It would be good to describe in the methods section or here in the results, how and why the results for diarrhoea prevalence in 1995 and 2001 were collapsed to give a value of 22%. While the total values of diarrhoea prevalence are summarized each period at the bottom of the table, it is very hard to believe that in both 1995 and 2001 the prevalence was identical (22.2%) 8. Results. Trends in diarrhoea prevalence. The investigators essentially describe the content of table 2 and do not add more detail from what I can see for myself in the table. There is therefore significant duplication between the narrative and table 2. 9. Results. Trends in diarrhoea prevalence. The investigators report trends in prevalence rates but do not comment whether the changes are statistically significant or give any confidence intervals around the percentage changes for readers to determine for themselves the significance of the changes. 10. Results. Table 2. While the MICS and DHS disaggregate results by Province, it is difficult to assess if the changes over time were significant as the sample sizes included in the surveys were so different. The investigators do not comment on this but report and interpret the percentage changes without any qualification. Several of the trends are probably not significant. 11. Results. Decomposition analysis. The investigators apply the decomposition analysis approach to data from individual Provinces. I am not sure if this is appropriate given the much small samples in these Provinces. The investigators should indicate the validity of applying the approach to the disaggregated data as well as to the Overall dataset. 12. It would be a much stronger analysis if changes in U5 mortality over the same period were included as part of the decomposition analysis. This would demonstrate the integrity and relevance of the results and also place a much higher value on interventions to improve environmental conditions. 13. Discussion. The first page of the discussion essentially repeats the first paragraph of the results section. It then goes on to discuss findings from other studies and what is thought to influence diarrhoeal disease based on findings from general studies. It would be much better to focus the discussion on the results of this analysis. From a style perspective, I would also suggest that the investigators do not use rhetorical questions to lead into an issue: <i>Did the environmental sanitation and What could explain?</i> 14. The rest of the Discussion could also benefit from additional editing as there is significant repetition with other parts of the manuscript. <p>Minor comments</p> <ol style="list-style-type: none"> 15. Introduction. References 1-3. There are more current estimates of the number of children dying from diarrhoea that could be included. For example, in 2010 it is estimated that about 740,000 children died of diarrhoea. This represents about 11% of all child mortality. WHO. Global Health Observatory (http://www.who.int/gho/child_health/en/index.html). In the
--	--

	<p>DRC about 13% of child deaths were diarrhoea related i.e. about 60,450 in 2010. The total child deaths globally have now been reported to have dropped further to 6.9m.</p> <p>16. Introduction. Suggest to rephrase '... environmental degradation ..' to deterioration in environmental conditions.</p> <p>17. Introduction. The effect of civil strife is referenced. However, it was not uniform across all of the DRC. Given that the data differs by Province, it would be good to include a brief comment about the Provinces that were most affected by the civil/military unrest.</p> <p>18. Data and methods. Variables. It would be good to qualify the statement that the '... vast majority of diarrhea is caused by infectious agents ...'. This is true of acute diarrhoea in low resource settings. However, this may not be true of persistent diarrhoea in wealthy countries.</p> <p>19. Data and methods. Variables. The description that 'Human faeces are the primary source ...' is not really Methods and should not be included under this section. While the three paragraphs are informative to some extent, they are not especially relevant to the Methods section and would be better placed in the Discussion or Background. Methods should be very focused on methods applied in data collection or analysis.</p> <p>20. Data and methods. Data. The description of the political events in the DRC are not particularly relevant to the methods section. The unrest within the DRC is mentioned in the Background and should not be duplicated here. Sufficient to state there was civil unrest or civil war.</p> <p>21. Data and methods. Data. There is no need to justify cluster sampling as a cost-saving measure as part of the methods section.</p>
--	---

REVIEWER	Parfait M Eloundou-Enyegue Cornell University
REVIEW RETURNED	25-Sep-2012

THE STUDY	A Y/N framework does not make it possible to give enough nuance. It would be better if reviewers had the opportunity to rate these items on a scale (say) from 0 to 5.
RESULTS & CONCLUSIONS	There seems to be a substantial inconsistency between the final recommendations and the main findings from the decomposition methods. The final conclusions may need some updating
GENERAL COMMENTS	<p>This paper is a straightforward assessment of recent trends in the prevalence of diarrheal disease in the DRC. It uses a massive set of data to explore this question and finds a noticeable decline that reflects not a compositional but a behavioral change. However, they are unsure about the extent to which their conclusions are affected by differences in the timing of data collection, hence the possibility of confounding by seasonality effects. The paper does address a very important topic in a relevant country and the issues it raises deserve attention but I would make several recommendations for improvement. These are given below in no special order.</p> <ul style="list-style-type: none"> • Title: as a title I would recommend the following : "Accounting for recent trends in the prevalence of diarrhea

	<p>in the DRC“</p> <ul style="list-style-type: none"> • General editing: I would recommend substantial editing for typos, grammar and style. The document is already very well written for two authors who are presumably not native speakers but the writing can be improved. • Design: the study design (p. 2) is better described as a repeated cross-section • Method: the decomposition approach used is a viable method to assess the historical changes during that time period, especially given that the authors, contrary to a basic decomposition, attempt to incorporate evidence from the initial fixed-effect regression analysis. I would however suggest the following: <ul style="list-style-type: none"> ○ Terminology: what the authors refer to as “performance” effect can be termed “behavioral” effect ○ Selection of independent variables: Maternal education or even family SES (given the economic transformations the authors refer to) would have been more salient policy variables in the analysis than age or province; if the authors stick to age, I would recommend breaking the categories down in terms of nursing, crawling, and walking ages, as the authors do towards the end of the paper; ○ Analysis of housing circumstances: in exploring the effects of housing conditions, the authors (table 3, page 11) contrast three categories: none; toilet or water, and toilet and water; I would have found it much more interesting to distinguish situations where they have toilet alone from those where there is water alone; furthermore, are there other housing circumstances (crowding, presence of other under-age children that deserve attention). These are better policy variables that would improve this analysis. • Method (ctd): I did not fully understand the unit of analysis and the focus of the fixed effects tables presented in Tables 5 and 6. The table is labeled “fixed effects of changes.” I would need a little more explanation about data construction for this particular analysis • Seasonality: Given the authors’ acknowledgment at the end that seasonality is likely to be a factor, is it possible to give the reader more background information about the periods of data collection for the surveys, and about what seasons are presumed to be more problematic (there is a paragraph to this effect at the end but it is unclear how long the surveys themselves lasted: the dates seem to be about the duration of the seasons themselves, not the duration of the survey). • Conclusions: the general conclusion and the recommendation that “improvement in household conditions “ as policy option (page 18) seems inconsistent with the
--	---

	decomposition findings (page 11). That deserves some explanation.
--	---

REVIEWER	Yazoume Ye, Senior Infectious Disease M&E Specialist ICF International, USA I know the authors and have published with them before.
REVIEW RETURNED	27-Sep-2012

GENERAL COMMENTS	<p>Review report.</p> <p>Article: Trends in the prevalence of diarrhea among children in the Democratic Republic of Congo (DRC). What is the source of changes?</p> <p>The author seeks to assess trends in diarrhea prevalence by age, province and household living conditions; and to identify the sources of variation in Democratic Republic of Congo (DRC). The data used for this analysis are from the MICS, and DHS.</p> <p>This is an interesting paper which addresses an important public health issues, particularly in country like the DRC. I read with interest the paper which needs some structuring specially the method and results section. The paper also will require an extensive editing as some sentences are very confusing. I have a number of observations that the authors might want to consider to improve the paper:</p> <ol style="list-style-type: none"> 1. The analysis is strong and well thought, however, the author should acknowledge the fact that the outcome (diarrhea) could be affected by measurement issue. It is based on self-reporting. While I admit that this is the best they could have at the population level, the authors failed to highlight this as limitations and discuss how this could have impacted on their results. The discussion section should be revised to address this issue. 2. It is not very clear to me if the change in living conditions contributed the decline in diarrhea prevalence. The text is contradictor. What is said in the abstract is different from what is said in the discussion- see line 40-42. The authors should reconcile and provide a clear take home message. 3. In the abstract the authors make reference to data quality that could have be responsible for the results observed. However failed to address this in the discussion. What are the specific issues related to data quality? In which way these could have affected the results? 4. I suggest deleting the first paragraph under section 1.1. These facts are known therefore is better to just provide references. 5. I suggest you delete the sentence: Cluster sampling is a cost-saving measure, without the need to list all the households. See section 1.2, Paragraph 2
-------------------------	---

	<p>6. I suggest the section 1.1 (Variable) comes after section 1.2- Data- For better flow. The third paragraph of section 1.2 will fit better under section 1.1</p> <p>7. Table 1 and its description will fit better under Result section as description of study sample. It should be the first subsection of that section.</p> <p>8. Section 2.1 Statistic methods- I suggest rephrasing the second sentence to. "Descriptive analysis assess the trends in diarrhea prevalence by exposure variables (child' age, province of residence, household living conditions)"</p> <p>9. The number of tables in the paper should be reduced to a maximum of 3. The authors should focus on the most important ones.</p>
--	--

VERSION 1 – AUTHOR RESPONSE

Reviewer 1:

Nigel Rollins, Scientist World health Organization, Switzerland

General

This paper examines the trends in diarrhea prevalence among children less than 5 years of age in regions of the Democratic Republic of Congo (DRC) using data from three population based surveys conducted in 1995, 2001 and 2007. The investigators present the background demographics of the Provinces surveyed and comment on the variation between the sample populations. Significant differences in diarrhea prevalence are evident over the three surveys.

The investigators use different analytical approaches to assess the causes of the changes over time.

The primary conclusion was that diarrhea remains a significant public health problem. The investigators infer that interventions to improve home conditions would reduce diarrhea prevalence.

While I commend the intention of the investigators to examine the available datasets to understand changes in diarrhea prevalence in the DRC, I do not feel the manuscript is strong enough yet to be published in the Journal. There is significant repetition and insufficient discussion of the data itself and the limitations of the approaches taken. There are also several grammatical points and style issues that could be improved.

Reply: We would like to thank the reviewer for his comments on our manuscript. We have now carefully checked the paper and remove repetitive sentences. The paper also has been revised by a colleague to collect any grammatical error throughout the manuscript.

Major comments

1. The investigators have used data collected in the 1995 and 2001 MICS Surveys and the 2007 DHS to determine relationships between patterns of diarrhea prevalence in children and causal factors.

The analyses and interpretation are therefore fundamentally limited by the quality of the survey data, which the investigators obviously cannot change. However, it is a critical limitation, which has important consequences for interpretation of the data. While the authors acknowledge this, they do not really examine the implications. For example, in the discussion, it is simply stated that 'there may also be some issues with data quality' but they do not really discuss how these issues might be resolved either with respect to the current analysis (they could corroborate findings by concurrently examining mortality data) or in the future e.g. what it means for sampling in future population surveys or design of questions.

Reply: The reviewer is right. Data quality has important consequences for the interpretation of our results. We have now elaborated on this issue as shown below (see 1st paragraph, page 14):

"Two hypotheses could explain the observed discrepancies between the living conditions and

changes in diarrhea prevalence in the DRC. Firstly, they may be some real improvement in health behavior, particularly the use of boiled drinking water and the adequate washing of hands after contact with adult and child stools. However, we cannot test such hypothesis because of lack of data. Secondly, there may be some issues with data quality. In conflicts situation, it might very difficult to collect reliable data. For instance, during the 2001 MICS, in the Eastern part, often interviewers had to stop their work and resume after several days. Rebels arrested a fieldworker for more than six weeks (MICS2) [8]. Furthermore, diarrhea prevalence is based on self-reporting. Mothers or caregivers can mis-declare diarrhea prevalence according to her understanding of diarrhea definition in the local language. Also, duration of data collection varies considerably by province (1 month in Kinshasa and about 2 months in Nord-Kivu and Katanga for the 2001 MICS). In addition, the distribution of children by some socioeconomic characteristics varies across survey. This is probably due to the use of an old national sampling frame from the 1984 Census. However, the methods used (decomposition and fixed effect regression models) control for variation in proportion during analyses. “

2. Introduction. As much as diarrhea burden is important, it would be good to include data on diarrhea deaths or all U5 deaths over the same intervals and to comment on the relation between diarrhea and subsequent health outcomes. The majority of diarrhea events are not associated with death and therefore it is important to contextualize the significance of the diarrhea burden.

Reply: This is a good suggestion from the reviewer. We have now linked estimates of childhood diarrhea and those from under-five mortality as suggested by the reviewer in the introduction (see the paragraph 2 of the introduction section, page 3).

3. Methods. 1.2. Data. The investigators comment that the sample designs and questionnaires of the three surveys are reported elsewhere. While the sampling approaches are reasonably understood, the exact questions used in each of the surveys to determine diarrhea morbidity are critical to the analysis. It would be relevant to outline the actual wording used in each survey to ascertain information on diarrhea prevalence and if/how acute diarrhea is differentiated from persistent diarrhea or dysentery.

Reply: We have inserted original questions in French used in each survey to determine diarrheal morbidity (See Box 1, page 5).

4. Methods 1.2. Data. The investigators describe how they define high quality water – are similar definitions used in MICS or DHS or are these definitions applied to specific categories of water supply that are included in the survey tools?

Reply: Both DHS and MICS used the same categorization. Following the reviewer’s query, we have listed categories of water supply and improved sanitation facility in the text as shown below (last paragraph on page 5):

“We define clean water or drinking water as water of sufficiently high quality that can be consumed or used with low risk of immediate or long-term harm. It is drawn from an improved drinking water sources protected from outside contamination, in particular from contamination with fecal matter including piped water (into residence or plot), public tap, tube well, protected dug wells and protected springs [,].”

5. Table 1. There are major regional variations in the number of children U5 included between Provinces over time. What were the reasons for the change in number of children included? If this is known it would be good to give some details in the methods section. Whatever the reason, the investigators need to comment on the implications.

Reply: We are not aware of the reasons of change in proportion of under-children per province. It’s likely that it was influenced by the sampling procedures used in different surveys. However, the methods used in this study (decomposition and fixed effect regression models) did not suggest that the proportion of children U5 influenced the observed decrease in diarrhea prevalence. In fact, the decomposition analysis was used to detect such influence. The text on the manuscript has been

revised to highlight this issue (see 1st paragraph page 9).

6. Statistical methods. The description of the decomposition approach is not very clear and leaves the reader (me) unclear about the interpretation. I also found it odd to include a justification of the method as part of the methods section. While it may be a standard approach in demographic analyses, it is not a method that I, as a clinician and public health scientist, have used and feel confident to interpret. The investigators state that the approach is simple and gives a clear understanding of trends. After reading the paper, I cannot say that I clearly understand the trends presented and did not find the application of the decomposition analysis particularly informative.

Reply: Thank for the comments. We have removed the justification of the statistical methods as suggested and improve the explanation of the decomposition approach (see page 6 and last paragraph page 10). Basically, the decomposition analysis allows assessing the historical changes during the considered time period. It divides the trends in child's diarrhea prevalence into change in population structure and change in improvement in health behavior or public health.

7. Results. It would be good to describe in the methods section or here in the results, how and why the results for diarrhea prevalence in 1995 and 2001 were collapsed to give a value of 22%. While the total values of diarrhea prevalence are summarized each period at the bottom of the table, it is very hard to believe that in both 1995 and 2001 the prevalence was identical (22.2%).

Reply: Current analyses exclude the 1995 MICS data for some methodological reasons (E.g. problem with the variable maternal education). Furthermore, there is little or no change between 1995 and 2001. Again, we are also surprised to see that there were no change in diarrhea prevalence between 1995 and 2001.

8. Results. Trends in diarrhea prevalence. The investigators essentially describe the content of table 2 and do not add more detail from what I can see for myself in the table. There is therefore significant duplication between the narrative and table 2.

Reply: We have removed the duplication in the current text and taken into account the comment in the revised version of the paper.

9. Results. Trends in diarrhea prevalence. The investigators report trends in prevalence rates but do not comment whether the changes are statistically significant or give any confidence intervals around the percentage changes for readers to determine for themselves the significance of the changes.

Reply: We have performed trend test using the Stata command "nptrend" that provide P-value for percentage changes (Table 2).

10. Results. Table 2. While the MICS and DHS disaggregate results by Province, it is difficult to assess if the changes over time were significant as the sample sizes included in the surveys were so different. The investigators do not comment on this but report and interpret the percentage changes without any qualification. Several of the trends are probably not significant.

Reply: We performed a formal statistical test of trend using the Stata command "nptrend" that provide P-value for percentage changes and most of the tests were statistically significant as shown in Table 2 (see Table page 9). Furthermore, the decomposition analysis and the Fixed Effect models performed allow to show whether changes in diarrhea prevalence is due to changes in proportion or not. In our case, variation of proportion and the sample size does not explain decrease in diarrhea prevalence.

11. Results. Decomposition analysis. The investigators apply the decomposition analysis approach to data from individual Provinces. I am not sure if this is appropriate given the much small samples in these Provinces. The investigators should indicate the validity of applying the approach to the disaggregated data as well as to the Overall dataset.

Reply: Perhaps, it was not clear in the previous version of the paper but now we have clarified the

method section. We did not apply the decomposition at the disaggregate level. Province of residence is an independent variable in the decomposition analysis. This assumes that diarrhea prevalence observed at the national level is the weighted average of prevalence observed in different categories of independent variables including province. This method is different from the Oaxaca Decomposition, which explains differences between groups at time t . The decomposition we used aims at assessing the historical changes between two dates. The Oaxaca Decomposition is applied at the individual level, whereas the current decomposition method is performed at the aggregate level. However, we have replaced province of residence and children's age from the analyses by maternal education and household wealth index as suggested by Prof. Parfait Eloundou, one of the reviewers and an expert in decomposition analysis we have performed.

12. It would be a much stronger analysis if changes in U5 mortality over the same period were included as part of the decomposition analysis. This would demonstrate the integrity and relevance of the results and also place a much higher value on interventions to improve environmental conditions.
Reply: Thanks for the suggestion. But the objectives of this paper were to: identify sources of changes in diarrhea prevalence in the DRC (changes in the distribution of children by living conditions and/or by maternal education) or changes in behavior; and to assess the contribution of each category in the observed change. However, the reviewer is right that an adjacent analysis of U5 mortality within the same period will strengthen the interventions strategies. As U5 in the DRC is higher, we think that analysis of changes in under-five mortality in the DRC should be done as a separate study in the future. For now, we have discussed trends in diarrhea prevalence in parallel to trends in under-five mortality in the DRC.

13. Discussion. The first page of the discussion essentially repeats the first paragraph of the results section. It then goes on to discuss findings from other studies and what is thought to influence diarrhea disease based on findings from general studies. It would be much better to focus the discussion on the results of this analysis. From a style perspective, I would also suggest that the investigators do not use rhetorical questions to lead into an issue: Did the environmental sanitation and... What could explain?

Reply: Thanks for the suggestion. The discussion section has been revised to take into account the reviewer's suggestions.

14. The rest of the Discussion could also benefit from additional editing as there is significant repetition with other parts of the manuscript.

Reply: Thanks for the suggestion. We have revised the text and removed the repetition.

Minor comments

15. Introduction. References 1-3. There are more current estimates of the number of children dying from diarrhea that could be included. For example, in 2010 it is estimated that about 740,000 children died of diarrhea. This represents about 11% of all child mortality. WHO. Global Health Observatory (http://www.who.int/gho/child_health/en/index.html). In the DRC about 13% of child deaths were diarrhea related i.e. about 60,450 in 2010. The total child deaths globally have now been reported to have dropped further to 6.9m.

Reply: Thanks for pointing to these new estimates which have been inserted in the text and referenced (see second paragraph of the introduction on page 3).

16. Introduction. Suggest to rephrase '... environmental degradation...' to deterioration in environmental conditions.

Reply: The introduction section has been revised accordingly.

17. Introduction. The effect of civil strife is referenced. However, it was not uniform across all of the DRC. Given that the data differs by Province, it would be good to include a brief comment about the

Provinces that were most affected by the civil/military unrest.

Reply: Thanks for the suggestion. A paragraph has been included in the data section as shown below (see page 4):

"This study uses two successive nationally representative household surveys: the 2001 MICS and the 2007 DHS. During the 2001 Multiple Indicators Cluster Survey (MICS) data collection from May 21 to August 28, 2001, 3 provinces were entirely under the control of the government (Kinshasa, Bas-Congo and Bandundu), 4 were partially administrated by rebels (Equateur, Katanga, Kasai-Oriental and Kasai Occidental), and 4 were entirely controlled by rebels (Oriental, Nord Kivu, Sud Kivu and Maniema). Though the 2007 DHS was carried out after the 2006 elections (February 2 to April 30, 2007 for Kinshasa, and from May 10 to August 31, 2007 for the remaining provinces), some villages and municipalities in the Eastern provinces of Nord-Kivu, Sud-Kivu and Oriental were under armed conflict."

18. Data and methods. Variables. It would be good to qualify the statement that the '... vast majority of diarrhea is caused by infectious agents ...'. This is true of acute diarrhea in low resource settings. However, this may not be true of persistent diarrhea in wealthy countries.

Reply: This sentence has been dropped (see Yazoume comments).

19. Data and methods. Variables. The description that 'Human faeces are the primary source ...' is not really Methods and should not be included under this section. While the three paragraphs are informative to some extent, they are not especially relevant to the Methods section and would be better placed in the Discussion or Background. Methods should be very focused on methods applied in data collection or analysis.

Reply: Thanks for the suggestion. The sentence has been removed from the method section.

20. Data and methods. Data. The description of the political events in the DRC are not particularly relevant to the methods section. The unrest within the DRC is mentioned in the Background and should not be duplicated here. Sufficient to state there was civil unrest or civil war.

Reply: Thanks for the suggestion. The suggested changes have been made.

21. Data and methods. Data. There is no need to justify cluster sampling as a cost-saving measure as part of the methods section.

Reply: Thanks for the suggestion. The sentence has been deleted.

Reviewer 2:

Parfait M Eloundou-Enyegue, Cornell University

There seems to be a substantial inconsistency between the final recommendations and the main findings from the decomposition methods. The final conclusions may need some updating

Reply: Thanks for the suggestion. We have updated the conclusion in line with our results.

This paper is a straightforward assessment of recent trends in the prevalence of diarrheal disease in the DRC. It uses a massive set of data to explore this question and finds a noticeable decline that reflects not a compositional but a behavioral change.

Reply: We thank the reviewer for his kind comments. Although, the method used is straightforward, the results have been proven to be insightful and informative in the DRC's context for public health policy making.

However, they are unsure about the extent to which their conclusions are affected by differences in the timing of data collection, hence the possibility of confounding by seasonality effects. The paper

does address a very important topic in a relevant country and the issues it raises deserve attention but I would make several recommendations for improvement. These are given below in no special order.

1. Title: as a title I would recommend the following: "Accounting for recent trends in the prevalence of diarrhea in the DRC"

Reply: The title has been changed as suggested to "Accounting for recent trends in the prevalence of diarrhea in the Democratic Republic of Congo (DRC)"

2. General editing: I would recommend substantial editing for typos, grammar and style. The document is already very well written for two authors who are presumably not native speakers but the writing can be improved.

Reply: Thank you for the suggestion. The manuscript has been read by a colleague to check for typos etc... and we hope that it has improved compared with the previous version.

3. Design: the study design (p. 2) is better described as a repeated cross-section

Reply: The study design has been changed as suggested. We prefer the term consecutive cross-sectional surveys.

4. Method: the decomposition approach used is a viable method to assess the historical changes during that time period, especially given that the authors, contrary to a basic decomposition, attempt to incorporate evidence from the initial fixed-effect regression analysis. I would however suggest the following:

4.1. Terminology: what the authors refer to as "performance" effect can be termed "behavioral" effect

Reply: Performance effect has been changed to behavioral effect as suggested throughout the text.

4.2. Selection of independent variables: Maternal education or even family SES (given the economic transformations the authors refer to) would have been more salient policy variables in the analysis than age or province; if the authors stick to age, I would recommend breaking the categories down in terms of nursing, crawling, and walking ages, as the authors do towards the end of the paper;

Reply: Independent variables include now access to water and sanitation, maternal education and household wealth index as suggested (see page 5).

4.3. Analysis of housing circumstances: in exploring the effects of housing conditions, the authors (table 3, page 11) contrast three categories: none; toilet or water, and toilet and water; I would have found it much more interesting to distinguish situations where they have toilet alone from those where there is water alone; furthermore, are there other housing circumstances (crowding, presence of other under-age children that deserve attention). These are better policy variables that would improve this analysis.

Reply: The variable housing conditions has been split into 4 categories: none, toilet alone, water alone, water and toilet. The variables "Number of household members" and "Number of under-five children living in the household" have not been included in the final analysis because of lack of statistical association with diarrhea prevalence in the exploratory analyses.

4.3. Method (ctd): I did not fully understand the unit of analysis and the focus of the fixed effects tables presented in Tables 5 and 6. The table is labeled "fixed effects of changes." I would need a little more explanation about data construction for this particular analysis.

Reply: We have inserted a paragraph describing construction of the three datasets (1 per independent variable) used for performing the "Fixed effect models" (see page 5).

4.4. Seasonality: Given the authors' acknowledgment at the end that seasonality is likely to be a factor, is it possible to give the reader more background information about the periods of data collection for the surveys, and about what seasons are presumed to be more problematic (there is a paragraph to this effect at the end but it is unclear how long the surveys themselves lasted: the dates seem to be about the duration of the seasons themselves, not the duration of the survey).

Reply: we have provided data collection dates in the method section as shown below (see page 4):

“This study uses two successive nationally representative household surveys: the 2001 MICS and the 2007 DHS. During the 2001 Multiple Indicators Cluster Survey (MICS) data collection from May 21 to August 28, 2001, 3 provinces were entirely under the control of the government (Kinshasa, Bas-Congo and Bandundu), 4 were partially administrated by rebels (Equateur, Katanga, Kasai-Oriental and Kasai Occidental), and 4 were entirely controlled by rebels (Oriental, Nord Kivu, Sud Kivu and Maniema). Though the 2007 DHS was carried out after the 2006 elections (February 2 to April 30, 2007 for Kinshasa, and from May 10 to August 31, 2007 for the remaining provinces), some villages and municipalities in the Eastern provinces of Nord-Kivu, Sud-Kivu and Oriental were under armed conflict. “

4.5. Conclusions: the general conclusion and the recommendation that “improvement in household conditions “as policy option (page 18) seems inconsistent with the decomposition findings (page 11). That deserves some explanation.

Reply: Thanks for the suggestion. We have revised our conclusion in line with our results.

Reviewer 3:

Yazoume Ye,

Senior Infectious Disease M&E Specialist ICF International, USA

I know the authors and have published with them before.

The author seeks to assess trends in diarrhea prevalence by age, province and household living conditions; and to identify the sources of variation in Democratic Republic of Congo (DRC). The data used for this analysis are from the MICS, and DHS.

This is an interesting paper which addresses an important public health issues, particularly in country like the DRC. I read with interest the paper, which needs some structuring specially the method and results section. The paper also will require an extensive editing as some sentences are very confusing. I have a number of observations that the authors might want to consider to improve the paper:

Reply: We thank the reviewer for his kind comments.

1. The analysis is strong and well thought, however, the author should acknowledge the fact that the outcome (diarrhea) could be affected by measurement issue. It is based on self-reporting. While I admit that this is the best they could have at the population level, the authors failed to highlight this as limitations and discuss how this could have impacted on their results. The discussion section should be revised to address this issue.

Reply: Thanks for the suggestion. We have mentioned this issue in the discussion section (see second paragraph page 14) as shown below:

“Secondly, there may be some issues with data quality. In conflicts situation, it might very difficult to collect reliable data. For instance, during the 2001 MICS, in the Eastern part, often interviewers had to stop their work and resume after several days. Rebels arrested a fieldworker for more than six weeks (MICS2) [8]. Furthermore, diarrhea prevalence is based on self-reporting. Mothers or caregivers can mis-declare diarrhea prevalence according to her understanding of diarrhea definition in the local language. Also, duration of data collection varies considerably by province (1 month in Kinshasa and about 2 months in Nord-Kivu and Katanga for the 2001 MICS). In addition, the distribution of children by some socioeconomic characteristics varies across survey. This is probably due to the use of an old national sampling frame from the 1984 Census. However, the methods used (decomposition and fixed effect regression models) control for variation in proportion during analyses. “

2. It is not very clear to me if the change in living conditions contributed the decline in diarrhea prevalence. The text is contradictor. What is said in the abstract is different from what is said in the discussion- see line 40-42. The authors should reconcile and provide a clear take home message.

Reply: The abstract and discussion have been amended as shown here. “Changes in behavior and/or

in public health policy seem to be the likely main source of the observed changes as there were no significant changes in diarrhea prevalence associated with variation of the population structure” (see results section of the abstract).

3. In the abstract the authors make reference to data quality that could have be responsible for the results observed. However failed to address this in the discussion. What are the specific issues related to data quality? In which way these could have affected the results?

Reply: We have explained in more details in the discussion the issue data quality that might have influenced our results (see the discussion section on page 14).

4. I suggest deleting the first paragraph under section 1.1. These facts are known therefore is better to just provide references.

Reply: The sentence has been deleted as suggested.

5. I suggest you delete the sentence: Cluster sampling is a cost-saving measure, without the need to list all the households. See section 1.2, Paragraph 2

Reply: The sentence has been deleted as suggested

6. I suggest the section 1.1 (Variable) comes after section 1.2- Data- For better flow. The third paragraph of section 1.2 will fit better under section 1.1

Reply: Variable section has been moved and the third paragraph of the data section has been moved to the variable section as suggested (see page 5).

7. Table 1 and its description will fit better under Result section as description of study sample. It should be the first subsection of that section.

Reply: Table 1 and its description have been moved under the result section as suggested.

8. Section 2.1 Statistic methods-I suggest rephrasing the second sentence to. “Descriptive analysis assess the trends in diarrhea prevalence by exposure variables (child’ age, province of residence, household living conditions)”

Reply: We have implemented the suggestion.

9. The number of tables in the paper should be reduced to a maximum of 3. The authors should focus on the most important ones.

Reply: The number of Tables has been reduced to 4.

We would like to thank the reviewers for thoughtful comments and suggestions. We truly appreciate your interest in our work. We believe that as a result of the review process our paper has greatly improved and hope that it is now acceptable for publication in BMJ Open.

VERSION 2 – REVIEW

REVIEWER	Yazoume Ye Senior Infectious Disease M&E, ICF International, USA I published with the authors
REVIEW RETURNED	04-Nov-2012

GENERAL COMMENTS	I reviewed the revised version of the manuscript and my concerned were appropriately addressed in the new version.
-------------------------	--